Nematotoxicity of Marasmius oreades Agglutinin (MOA) Depends on Glycolipid Binding and Cysteine Protease Activity

Wohlschlager, T ; Butschi, A ; Zurfluh, K ; Vonesch, S C ; Auf dem Keller, U ; Gehrig, P ; Bleuler-Martinez, S ; Hengartner, M O ; Aebi, M ; Künzler, M

Abstract: Fruiting body lectins have been proposed to act as effector proteins in the defense of fungi against parasites and predators. The Marasmius oreades agglutinin (MOA) is a Gal 1,3Gal/GalNAc-specific lectin from the fairy ring mushroom that consists of an N-terminal ricin B-type lectin domain and a C-terminal dimerization domain. The latter domain shows structural similarity to catalytically active proteins, suggesting that, in addition to its carbohydrate-binding activity, MOA has an enzymatic function. Here, we demonstrate toxicity of MOA toward the model nematode Caenorhabditis elegans. This toxicity depends on binding of MOA to glycosphingolipids of the worm via its lectin domain. We show further that MOA has cysteine protease activity and demonstrate a critical role of this catalytic function in MOA-mediated nematotoxicity. The proteolytic activity of MOA was dependent on high Ca(2+) concentrations and favored by slightly alkaline pH, suggesting that these conditions trigger activation of the toxin at the target location. Our results suggest that MOA is a fungal toxin with intriguing similarities to bacterial binary toxins and has a protective function against fungivorous soil nematodes.

DOI: https://doi.org/10.1074/jbc.M111.258202

Originally published at:
Wohlschlager, T; Butschi, A; Zurfluh, K; Vonesch, S C; Auf dem Keller, U; Gehrig, P; Bleuler-Martinez, S; Hengartner, M O; Aebi, M; Künzler, M (2011). Nematotoxicity of Marasmius oreades Agglutinin (MOA) Depends on Glycolipid Binding and Cysteine Protease Activity. Journal of Biological Chemistry, 286(35):30337-30343.
DOI: https://doi.org/10.1074/jbc.M111.258202

JBC Papers in Press. Published on July 8, 2011 as Manuscript M111.258202
The latest version is at http://www.jbc.org/cgi/doi/10.1074/jbc.M111.258202

NEMATOTOXICITY OF *MARASMIUS OREADES* AGGLUTININ (MOA) DEPENDS ON GLYCOLIPID-BINDING AND CYSTEINE PROTEASE ACTIVITY

Therese Wohlschlager[1]*, Alex Butschi[2]*, Katrin Zurfluh[1], Sibylle C. Vonesch[2], Ulrich auf dem Keller[3], Peter Gehrig[4], Silvia Bleuler-Martinez[1], Michael O. Hengartner[2], Markus Aebi[1], and Markus Künzler[1]

*authors contributed equally to this work

1. Institute of Microbiology, ETH Zürich, CH-8093, Zürich, Switzerland
2. Institute of Molecular Life Sciences, University of Zürich, CH-8057, Zürich, Switzerland
3. Institute of Cell Biology, ETH Zürich, CH-8093, Zürich, Switzerland
4. Functional Genomics Center, University and ETH Zürich, CH-8057, Zürich, Switzerland

Running title: Nematotoxicity of the fungal chimerolectin MOA

Corresponding author: Markus Künzler, Wolfgang-Pauli-Str. 10, HCI F413, CH-8093 Zürich, Switzerland; Tel: +41 446324925; Fax: +41 446321148; e-mail: markus.kuenzler@micro.biol.ethz.ch

Fruiting body lectins have been proposed to act as effector proteins in the defense of fungi against parasites and predators. The *Marasmius oreades* agglutinin (MOA) is a Galα1,3Gal/GalNAc-specific lectin from the fairy ring mushroom that consists of an N-terminal, ricin B-type lectin domain and a C-terminal dimerization domain. Latter domain shows structural similarity to catalytically active proteins, suggesting that, in addition to its carbohydrate-binding activity, MOA has an enzymatic function. Here, we demonstrate toxicity of MOA towards the model nematode *Caenorhabditis elegans*. This toxicity depends on binding of MOA to glycosphingolipids of the worm via its lectin domain. We further show that MOA has cysteine protease activity and demonstrate a critical role of this catalytic function for MOA-mediated nematotoxicity. The proteolytic activity of MOA was dependent on high Ca^{2+} concentrations and favoured by slightly alkaline pH, suggesting that these conditions trigger activation of the toxin at the target location. Our results suggest that MOA is a fungal toxin with intriguing similarities to bacterial binary toxins and has a protective function against fungivorous soil nematodes.

A wide variety of lectins with different structures and specificities has been isolated from fungi (1). Recently, cytoplasmic fungal lectins, also referred to as fruiting body lectins, were proposed to be part of a fungal defense system against parasites and predators. This hypothesis is based on the carbohydrate-binding dependent toxicity of several fungal lectins towards nematodes, insect larvae and amoebae (2). The exclusive and abundant expression of these lectins in fruiting bodies or sclerotia of the host fungus may result in a special protection of these reproductive organs. It is proposed that predators such as fungivorous nematodes ingest the cytoplasmic lectin when feeding on the content of the fungal cell and that binding of the lectin to specific glycans in the intestine of the worm induces toxicity by an unknown mechanism (2-5).

The *Marasmius oreades* agglutinin (MOA) is a type B erythrocyte agglutinating lectin that was isolated from fruiting bodies of the fairy ring mushroom. The dimeric lectin consists of 293-residue protomers and has been shown to specifically recognize Galα1,3-containing structures, exhibiting highest affinity for the branched blood group B trisaccharide Galα1,3(Fucα1,2)Gal (6). Crystal structures of MOA in complex with the xenotransplantation epitope and the blood group B trisaccharide revealed an N-terminal ricin B-type lectin domain with three canonical carbohydrate binding sites and a C-terminal domain that is involved in dimerization. The C-terminal domain showed structural homology to catalytically active proteins such as peptide-N-glycanases, N-acetyltransferases, transglutaminases and cysteine proteases. This structural similarity was limited to a putative catalytic triade consisting of Cys215, His257 and Glu274 (7,8). However, no catalytic activity of MOA has been reported so far. Moreover, the biological target glycan and the function of the lectin remained unclear. Like other fruiting body lectins, MOA lacks a classical secretion signal and is thus predicted to be localized in the cytoplasm. An endogenous function of the lectin is therefore unlikely as glycosylated structures are generally secreted or localized in the extracellular space (9). Accordingly, no glycoconjugate recognized by MOA has thus far been identified in fungi.

In this study, we investigated the possible role of MOA as a fungal defense molecule against nematodes, using *C. elegans* as a genetically tractable model system. First, we demonstrate carbohydrate-binding dependent nematotoxicity of MOA and identify glycosphingolipids as the ligands of MOA in *C. elegans*. Second, we show that MOA-mediated toxicity is dependent on the previously suggested catalytic function that we identified as Ca^{2+} dependent cysteine protease activity. Our results suggest that MOA is a fungal toxin that protects the fruiting body from predators such as fungivorous nematodes.

EXPERIMENTAL PROCEDURES

Strains, cultivation conditions, cloning and site-directed mutagenesis. Strains, primer sequences and experimental procedures can be found in *SI Materials and Methods*.

C. elegans toxicity assays and statistical analysis. Biotoxicity assays with *C. elegans* were performed as described (10). Results were analyzed using a t-test for pairwise comparisons.

Protein expression, purification and biotinylation. MOA and MOA(C215A) were expressed recombinantly in *E. coli* and purified as described (6,11). Detailed information can be found in *SI Materials and Methods*. t-Cry5B was prepared as previously described (12). For TLC overlay analysis, purified MOA and t-Cry5B were labeled with EZ-Link Sulfo-NHS-Biotin (Pierce) using a 20-fold molar excess of labeling reagent. Labeling proceeded for two hours on ice followed by desalting on a PD-10 column.

Screen for MOA-resistant C. elegans mutants using Mos1 insertional mutagenesis and light microscopy of C. elegans. Mos1 insertional mutagenesis, stereomicroscopy and differential interference contrast (DIC) microscopy of *C. elegans* was performed as described (3).

Preparation and analysis of glycolipids from C. elegans. Glycolipid extraction from *C. elegans* N2, *bre-2* and *bre-4* strains and TLC overlay staining was performed as described by Barrows *et al* (13). For overlay analysis biotinylated MOA and t-Cry5B were used at a concentration of 150nM.

In vitro protease assay. Briefly, 20 µg of denatured RNase A was incubated with 0.1 ng to 1 µg of purified recombinant MOA or MOA(C215A) in assay buffer (40 mM Tris-HCl pH 8.0, 1 mM $CaCl_2$, 1% NP-40) in a final reaction volume of 30 µl for 1 h at 37 °C. Samples were analyzed by SDS-PAGE and Coomassie blue staining. Detailed information is provided in *SI Materials and Methods*.

Cleavage site specificity. 40 µg of bovine asialofetuin, bovine casein and human hemoglobin were digested with 100 ng of MOA in 40mM Tris-HCl pH8.0, 1mM $CaCl_2$ for 1h at 37°C. The reaction was stopped by addition of 25 mM EDTA. Samples were desalted using C18 ZipTips (Millipore) and analyzed on an LTQ-Orbitrap XL mass spectrometer (Thermo Fischer Scientific, Bremen, Germany) coupled to an Eksigent-Nano-HPLC system (Eksigent Technologies, Dublin (CA), USA). Detailed information on MS analysis and data evaluation is provided in *SI Materials and Methods*.

RESULTS

MOA inhibits C. elegans development and amoebal growth dependent on its carbohydrate-binding ability. Based on the results with other fruiting body lectins (2), we tested MOA for toxicity towards the nematode *Caenorhabditis elegans* and the amoeba *Acanthamoeba castellanii*. These biotoxicity assays involved feeding of the organisms with MOA-expressing *E. coli* as described previously (10) and revealed a clear toxicity of MOA towards *C. elegans* and *A. castellanii* (Fig. 1; Fig. S1). Both *C. elegans* and *A. castellanii* were fed with *E. coli* cells expressing either authentic MOA, the C-terminal deletion mutant MOA(ΔC) or the carbohydrate-binding deficient variant MOA(Q46A,W138A) (7). In a control experiment, test organisms were grown on vector-containing *E. coli* transformants. In case of *C. elegans*, severe inhibition of larval development was observed for N2 wildtype and *pmk-1(km25)* mutant animals feeding on MOA-expressing bacteria compared to the vector control ($p<0.05$) (Fig. 1*A*). The *pmk-1(km25)* mutation in the p38 mitogen-activated protein kinase (MAPK) pathway was previously shown to cause hypersensitivity to different kinds of stress including exposure to fungal lectins (3). In contrast, MOA(Q46A,W138A) had no significant effect on the development of the worms ($p>0.05$). Similar results were obtained with *A. castellanii* ($p<0.05$) (Fig. S1). Previous studies showed that residues Q46 and W138 are involved in glycan coordination (7,8) and that the binding activity of the double mutant to

thyroglobulin is reduced compared to wild-type MOA (14). The lacking toxicity of MOA(Q46A,W138A) suggests that nematotoxicity of MOA is dependent on the carbohydrate-binding activity of the protein. This is also supported by the decreased nematotoxicity of MOA(ΔC) ($p<0.05$), a mutant that showed reduced carbohydrate-binding activity possibly due to loss of dimerization (14) (Fig. 1*A*).

Physiological effects of intoxication were studied by comparative light- and differential interference contrast (DIC) microscopy of *C. elegans* L4 larvae fed with MOA-expressing or vector-containing *E. coli* cells. MOA-intoxicated animals showed an expanded lumen of the anterior intestine compared to control animals (Fig. 1*B*). A similar phenotype was observed with other nematotoxic fungal lectins (3) and bacterial toxins (15).

The ligand of MOA in C. elegans is a glycosphingolipid. In order to identify the target glycoconjugate of MOA in *C. elegans*, a forward genetic screen for mutations conferring MOA-resistance was performed. Strain *pmk-1(km25)* was used to increase the sensitivity of the screen. Four MOA-resistant mutants were isolated in the screen, all of which carried mutations in exons of the *bre-3* gene. *Bre-3(op504)* and *bre-3(op505)* mutants had a *Mos1* insertion in the seventh exon of *bre-3*. *Bre-3(op506)* mutants carried a *Mos1* insertion in the fifth exon of the same gene. In an independent screen, one mutant *(op508)* with a *Mos1* insertion site identical to the one of *bre-3(op504)* was identified (Fig. 2*A*). *Bre-3* encodes for a glycosyltransferase that acts in a biosynthetic pathway building up an invertebrate-specific glycosphingolipid that is recognized by the *Bacillus thuringiensis* crystal toxin Cry5B (16). The oligosaccharide core of the ceramide-linked glycan is synthesized by the sequential action of BRE-3 (β1,4-mannosyltransferase), BRE-5 (β1,3-GlcNAc-transferase), BRE-4 (β1,4-GalNAc-transferase) and BRE-2 (putative β1,3-galactosyltransferase). *Bre-1* encodes for an enzyme involved in the conversion of GDP-mannose into GDP-fucose (17). Given the resistance of *bre-3* mutant animals we tested worms with mutations in the other *bre*-genes for sensitivity towards MOA. We observed that also mutations in *bre-4(ye27)* and *bre-5(ye17)* genes conferred resistance to MOA ($p<0.05$) whereas *bre-1(ye4)* and *bre-2(ye31)* mutant worms were sensitive to intoxication with MOA ($p>0.05$) (Fig. 2*B*). These results suggest that fungal MOA, like bacterial Cry5B, targets a glycosphingolipid but that the species is different from the one recognized by Cry5B.

In order to confirm the genetic data, direct binding of MOA to glycolipids isolated from N2, *bre-2(ye31),* and *bre-4(ye27)* mutant worms was assessed in a thin layer chromatography (TLC) overlay. For this purpose lipids were extracted from *C. elegans*, partitioned into lower phase lipids and upper phase glycolipids and subsequently purified as described by Barrows *et al*. (13). Purified glycolipids were separated by TLC using chloroform:methanol:water (4:4:1) as a solvent and visualized using Orcinol sulfate (Fig. 2*C*, lane 2-4). The main glycolipid bands were observed in positions similar to those reported previously (12,16). TLC overlay staining of upper phase glycolipids with MOA revealed binding to N2 and *bre-2(ye31)* but not to *bre-4(ye27)* glycolipids (Fig. 2*C,* lane 5-7). This is consistent with MOA-susceptibility of N2 and *bre-2(ye31)* worms and resistance of *bre-4(ye27)* animals. To identify the glycolipid species that is recognized by MOA, comparative overlay staining with truncated Cry5B (t-Cry5B) as described by Ideo *et al*. (12) was performed. The pattern of the Orcinol staining and the t-Cry5B overlay of N2- derived glycolipids (Fig. 2*C*, lane 1) indicated that the main glycolipid species recognized by MOA corresponds to component D according to the classification of Griffitts *et al.* that has been previously characterized (16,18). Binding to the glycan part of this glycosphingolipid of the arthroseries (Galα1,3GalNAcβ1,4GlcNAcβ1,3Manβ1,4Glc, Fig. 2*D*) (16,18) is in agreement with the Galα1,3Gal specificity of MOA. Finally, MS and MS/MS analysis of glycolipids extracted from the main TLC band recognized by MOA further supported the presence of component D including the ceramide moiety composed of a saturated hydroxy fatty acid with 22 and 24 carbon atoms, respectively (Fig. S2). Besides component D, MOA is able to bind to a number of unidentified glycolipid species from both N2 and *bre-2(ye31)* animals (Fig. 2*C*, lanes 5, 6). These glycolipid species are absent from the *bre-4(ye27)* animals.

MOA shows in vitro protease activity. X-ray crystallographic studies of MOA in complex with the xenotransplantation epitope and the blood group B trisaccharide revealed that the C-

terminal part of MOA adopts a fold that is observed in catalytically active proteins such as peptide N-glycanases, N-acetyltransferases, transglutaminases and cysteine proteases (7,8). As the highest homology was observed to a peptide N-glycanase (PNGase) we tested purified recombinant MOA in an *in vitro* PNGase assay described by Suzuki (19). However, instead of deglycosylation of the glycoprotein RNase B, degradation of the substrate was observed, indicating that MOA may have protease- rather than peptide N-glycanase activity. To confirm this hypothesis, we performed an *in vitro* protease assay with non-glycosylated RNase A as substrate (Fig. 3). MOA efficiently degraded denatured RNase A within 1 h at 37 °C. In contrast, MOA(C215A), in which the cysteine residue of the putative catalytic triad was replaced by an alanine, had no effect on RNase A, confirming that the observed protease activity was MOA-dependent. In accordance with these results, degradation of the substrate protein by MOA was abolished in the presence of 50 µM E-64, an irreversible, highly selective cysteine protease inhibitor, but not in the presence of 2 mM PMSF, an irreversible serine protease inhibitor or 2 mM $AgNO_3$, a potent trypsin and chymotrypsin inhibitor (Fig. 3*A*). Based on these results, we conclude that MOA has, in addition to its glycan-binding-properties, cysteine protease activity.

Structural homology to catalytically active proteins was first observed in a crystal structure of MOA in complex with the xenotransplantation epitope (7). A second crystal structure of MOA complexed with the blood group B trisaccharide revealed that binding of two Ca^{2+} ions close to the hypothetical active site in the C-terminal domain causes a conformational change opening up a large cleft containing the potential active site (8). To investigate the influence of divalent cations on MOA protease activity, reactions in assay buffer supplemented with 1mM $CaCl_2$, $MgCl_2$, $MnCl_2$, and $ZnCl_2$, respectively, were performed. In the presence of Ca^{2+} ions proteolytic degradation of RNase A by MOA was observed. In contrast, MOA showed hardly any activity in buffers supplemented with $MgCl_2$, $MnCl_2$ or $ZnCl_2$ or in the absence of divalent cations (Fig. 3*B*). This observation is in accordance with the structural data and suggests that Ca^{2+} binding activates the catalytic activity of MOA. We found that a minimal concentration of 50 µM was required for MOA to exhibit protease activity and that this activity further increased upon raising the Ca^{2+} concentration up to 5 mM (Fig. S3).

To test the influence of the pH on the reaction, protease assays in buffers of different pH-values were performed. Complete degradation of RNase A by MOA was observed at pH 8.0. At higher pH, MOA only partially degraded the substrate whereas under acidic conditions the enzyme was not active (Fig. 3*C*). Therefore MOA protease activity is favoured by a slightly alkaline pH. Presence or absence of the carbohydrate ligand Galα1,3Galβ1,4GlcNAc did not affect the proteolytic activity of MOA (Fig. S4).

In order to characterize the substrate specificity of MOA, we tested its activity towards different denatured and native protein substrates in the *in vitro* assay. Bovine serum albumin (BSA), fluorescein isothiocyanate (FITC) labeled casein, asialofetuin, and hemoglobin were degraded when denatured before incubation with MOA (Fig. S5*A*). In the native state, only FITC-casein, asialofetuin and hemoglobin were susceptible to MOA protease activity whereas BSA, RNase A and RNase B were not degraded (Fig. S5*B*). Proteolytic degradation of FITC-casein by MOA could also be detected in a fluorometric assay, where it showed a 3.6 fold higher activity/mole at pH 8.0 than the cysteine protease papain at pH 7.5 (Fig. S6). As MOA did not act on every tested native protein substrate we conclude that the protease has a certain substrate specificity.

To investigate the cleavage site specificity of MOA, native asialofetuin, casein or hemoglobin were digested with purified MOA and the resulting peptides were analyzed by LC-ESI-MS/MS. This procedure yielded 83 cleavage sites that were used to create an iceLogo (20) of the preferred sequence specificity, where the cleavage occurs between position 4 and 5 (Fig. 3*D*). The results of this analysis suggest that MOA prefers proline and proline or valine at position 0 and 2, respectively. A digest of *C. elegans* proteome with MOA yielded similar results (data not shown), confirming the specificity observed for the model substrates.

Protease activity of MOA is required for nematotoxicity. To test whether the protease activity of MOA plays a role in toxicity towards *C. elegans*, worms were fed with *E. coli* expressing the catalytic site mutant MOA(C215A). This single site mutant forms

dimers like wild-type MOA as confirmed by analytical gel filtration (Fig. S7). The toxicity of MOA(C215A) was significantly reduced in N2 animals and *pmk-1(km25)* mutants compared to wild-type MOA ($p<0.05$) (Fig. 4*A*). Similar results were obtained for *A. castellanii* ($p<0.05$) (Fig. S1). Based on these results, we conclude that, in addition to carbohydrate-binding, the cysteine protease activity of MOA is necessary for full toxicity. In *C. elegans*, the contribution of the catalytic domain to the toxicity of MOA is also apparent by the fact that *pmk-1(km25)* mutant and N2 wild type worms are equally susceptible towards wild-type MOA (Fig 1*A*, Fig 4*A*). This is in contrast to nematotoxic fungal lectins lacking a catalytic domain which display a higher toxicity to *pmk-1(km25)* mutant worms (2,3). Accordingly, abolishment of the catalytic activity, as in the case of MOA(C215A) and MOA(ΔC), results in higher susceptibility of *pmk-1(km25)* worms compared to N2 worms ((Fig 1*A*, Fig 4*A*). The higher toxicity of MOA(C215A) in comparison to MOA(ΔC) can be explained by the lack of dimerization in latter case. These results suggest that the toxicity mechanism of MOA differs from the one of nematotoxic lectins without a catalytic domain.

As the protease activity of MOA could be inhibited by E-64 *in vitro*, *pmk-1(km25)* animals were grown on MOA expressing bacteria in the presence of 10 to 500 µM of this highly selective cysteine protease inhibitor. The cell-permeable inhibitor significantly reduced MOA-mediated toxicity in a dose-dependent manner. On the other hand, E-64 did not impair development of worms feeding on vector-containing *E. coli* (Fig. 4*B*). These results confirm the role of MOA cysteine protease activity in nematotoxicity. As observed for the catalytic site mutant in the case of *pmk-1(km25)* worms, toxicity was not completely abolished by inhibiting protease activity even at the highest concentration of E-64. Remaining toxicity at inhibitor concentrations higher than 50 µM can be referred to the toxicity mediated by the lectin domains of the dimer, which are not affected by E-64.

DISCUSSION

Fruiting body lectins have been proposed to be part of the fungal defense system against predators and parasites (2). In agreement with such a function, we demonstrate that the fruiting body lectin MOA is toxic towards the nematode *C. elegans* and the amoeba *A. castellanii*. As MOA mutants deficient in carbohydrate-binding exhibited no or significantly reduced toxicity for both organisms, we conclude that the lectin-activity of the protein is essential for its toxicity. Morphological changes of MOA-intoxicated worms include an enlargement of the intestinal lumen, an effect that has previously been reported for the fungal lectin CGL2 from *C. cinerea* (3) and the *B. thuringiensis* crystal toxin Cry5B (15). These results suggest that binding of the lectin to a target glycan in the worm intestinal epithelium is essential for nematotoxicity. We hypothesize that the lectin is stored in the cytoplasm of the fungal cell and is released and transferred to the intestine of fungivorous nematodes upon predation.

Using *C. elegans* genetics, we could identify the target glycoconjugate of MOA in this organism that is responsible for its toxicity. Similar to the nematotoxic Cry5B toxin from *Bacillus thuringiensis*, MOA binds to different species of the invertebrate-specific arthroseries of glycosphingolipids in *C. elegans*. Based on TLC overlays and MS analysis, the main species recognized by MOA is component D according to the annotation by Griffitts *et al.* (16). This finding is consistent with the genetic results as *bre-3*, *bre-5* and *bre-4* (but not *bre-2*) encode glycosyltransferases involved in the biosynthesis of this specific glycosphingolipid. In addition to this previously identified glycosphingolipid species, MOA bound to several unknown glycolipids present in N2 and *bre-2(ye31)* but not in the *bre-4(ye27)* mutant worms. These results indicate the presence of multiple additional glycosphingolipid species carrying terminal Galα1,3Gal in *C. elegans*. The full sensitivity of *bre-1(ye4)* worms suggests that in contrast to glycosphingolipid-binding Cry5B, fucosylation of the lipid is not required for MOA binding.

The putative catalytic activity of MOA was investigated in *in vitro* assays with purified recombinant protein. In these assays, MOA showed proteolytic activity towards several native and denatured model substrates. The lack of activity towards some native protein substrates suggests a certain substrate specificity. The proteolytic activity of MOA was dependent on a cysteine residue in the predicted catalytic site, on the presence of Ca^{2+} and had a pH optimum of 8.0. The C-terminal domain of

MOA does not show homology to any existing proteinase family currently classified in the MEROPS database. Interestingly, the peptidase domain of MOA is the smallest in terms of residues that has ever been described and is even smaller than the HIV1 retropepsin. The crystal structure suggests that the active site cysteine is very close to the calcium binding sites, and the glutamine that would be an ideal candidate for forming the oxyanion hole is also a calcium ligand. Thus, the chemistry around the active site of this protease domain appears to be very 'busy' (N. D. Rawlings, personal communication).

Intriguingly, we could show that, in addition to carbohydrate-binding, the proteolytic activity of MOA is required for full toxicity towards *C. elegans* as toxicity of a catalytic site mutant is decreased and toxic effects of MOA are reduced in the presence of the highly specific cysteine protease inhibitor E-64. The equal susceptibility of N2 and *pmk-1(km25)* worms towards wild type MOA, as opposed to catalytically inactive MOA variants, suggests that the catalytic activity adds an additional twist to this lectin with regard to its mechanism of toxicity.

As determined by *in vitro* digests of different model substrates, MOA preferentially cleaves sequences with proline or valine at position P2. Interestingly, a similar sequence specificity has been described for calpains, a family of calcium-dependent, non-lysosomal cysteine proteases that are expressed ubiquitously in mammals and many other organisms (21) and are involved in regulation of signal transduction pathways, cell motility, and apoptosis (22,23). In contrast to calpain that proteolyzes its substrates in a limited manner exhibiting modulatory functions by cutting interdomain regions, our analysis using model substrates suggests that MOA digests protein substrates to small peptides. The specificity for sequences containing proline may allow MOA to act on a wide range of proteins, as proline residues are generally located in turns at the protein surface. However, further studies with more physiological substrate proteins are required to strengthen this hypothesis.

Chimeric lectins homologous to MOA are present in the basidiomycetes *Polyporus squamosus* (PSL) (24) and *Schizophyllum commune* (SCA) (25). In PSL, the catalytic triad and the four Asp residues taking part in Ca^{2+}-binding are conserved. The lectin domain of PSL, however, has been shown to be specific for Neu5Acα2,6Galβ1,4Glc/GlcNAc groups of N-glycans (24,26,27). In contrast, SCA has a similar carbohydrate-binding specificity as MOA (will be published elsewhere) but the catalytic residues Cys, His and Glu of MOA are replaced by Ser, Ser and Asp, respectively, and only one of the four Asp residues involved in Ca^{2+}-binding is conserved suggesting that the cysteine protease activity is not conserved in this homologue.

As one possible toxicity mechanism, binding of MOA to glycosphingolipids in the nematode intestinal epithelium may result in digestion of surrounding proteins and ultimately lead to disintegration of the intestinal epithelium. A similar mechanism has been proposed for the maize cysteine protease Mir1-CP, which has been shown to impair the growth of fall armyworm (*Spodoptera frugiperda*) larvae by disrupting the insect's peritrophic matrix in the intestine (28-30). Such proteases are used by plants, besides lectins and protease inhibitors, as defense molecules against herbivores (31) in line with our hypothesis of MOA being involved in fungal defense. Interestingly, Mir1-CP was recently shown to synergize the insecticidal activity of *Bacillus thuringiensis* toxins (32).

On the other hand, the carbohydrate-binding deficient but protease-proficient variant MOA(Q46A,W138A) was absolutely non-toxic even at the high lectin concentrations in the biotoxicity assay with recombinant bacteria (Fig. 1) and the lumen of the *C. elegans* intestine is thought to be slightly acidic based on the pH optima of digestive hydrolases (33), arguing against such a mechanism. In addition, the nature of its ligand (receptor) as well as the structural and functional organization of MOA are strikingly similar to bacterial AB_5 toxins such as the Cholera and the Shiga toxin. These proteins consist of a non-toxic B-subunit which binds to specific glycolipids and an enzymatically active A-subunit that induces toxicity by blocking the GTPase-activity of G_α and activating adenylate cyclase (Cholera toxin) or by inactivating ribosomal protein biosynthesis (Shiga toxin). The intoxication process by these proteins is initiated by attachment of the toxin via the B-subunit followed by internalization of the AB_5-receptor complex by receptor mediated endocytosis and retrograde transport to the endoplasmic reticulum (ER) (34). Finally, the A-subunit of these toxins is translocated to the cytoplasm by the sec translocon. This final step is omitted in the case of the subtilase cytotoxin

SubAB that induces cell death by exhibiting highly specific protease activity towards the ER-resident chaperone BiP (35). It is remarkable that SubAB produces pathological features in mice that strongly resemble the hemolytic uraemic syndrome (HUS) in humans (35), a characteristic that has also been observed for MOA (36). These results suggest that SubAB and MOA may have a similar toxicity mechanism. In accordance with such a hypothesis, the combination of neutral pH (37) and a Ca^{2+} concentration of more than 100 μM in the endoplasmic reticulum (38) meet the requirements of MOA to exhibit protease activity. On the other hand, MOA is not expected to be catalytically active at the neutral pH but low Ca^{2+} concentration (~ 100 nM) of the cytosol (37,39) which allows storage of this toxic protein in an inactive form in the cytoplasm of the fungal cell. Similarly, the slightly acidic pH in the intestinal lumen of the nematode should inhibit the proteolytic activity of MOA despite of the presumably high Ca^{2+} concentrations in this compartment (33). Such a regulation of its proteolytic activity via pH and Ca^{2+} concentration could explain the apparent lack of autoproteolytic activation or of pronounced substrate specificity of MOA. We therefore hypothesize that, upon binding to the glycosphingolipid receptor present on intestinal epithelial cells in the nematode, MOA is internalized to the early endosomal compartment by endocytosis and transported *via* the Golgi to the ER, where its catalytic domain is activated and may induce toxicity by degrading proteins that are in the folding process (Fig. S8). Further experiments are required to corroborate this hypothetical toxicity mechanism.

REFERENCES

1. Goldstein, I. J., Winter, H. C. (2007) *Mushroom Lectins.*, Elsevier, Oxford
2. Bleuler-Martinez, S., Butschi, A., Garbani, M., Walti, M. A., Wohlschlager, T., Potthoff, E., Sabotic, J., Pohleven, J., Luthy, P., Hengartner, M. O., Aebi, M., and Kunzler, M. (2011) *Mol Ecol*
3. Butschi, A., Titz, A., Wälti, M. A., Olieric, V., Paschinger, K., Nöbauer, K., Guo, X., Seeberger, P. H., Wilson, I. B. H., Aebi, M., Hengartner, M. O., and Künzler, M. (2009) *PLoS Pathog* **6**, e1000717
4. Trigueros, V., Lougarre, A., Ali-Ahmed, D., Rahbe, Y., Guillot, J., Chavant, L., Fournier, D., and Paquereau, L. (2003) *Biochim Biophys Acta* **1621**, 292-298
5. Hamshou M, S. G., Shahidi-Noghabi S, De Geyter E, Lannoo N, Van Damme EJ. (2010) *Insect Biochem Mol Biol.* **40**, 883-890
6. Winter, H. C., Mostafapour, K., and Goldstein, I. J. (2002) *J Biol Chem* **277**, 14996-15001
7. Grahn, E., Askarieh, G., Holmner, A., Tateno, H., Winter, H. C., Goldstein, I. J., and Krengell, U. (2007) *Journal of Molecular Biology* **369**, 710-721
8. Grahn, E. M., Winter, H. C., Tateno, H., Goldstein, I. J., and Krengel, U. (2009) *J Mol Biol* **390**, 457-466
9. Varki, A., Esko, JD, Colley, KJ. (2009) Cellular Organization of Glycosylation. . in *Essentials of Glycobiology.* (Varki, A., Cummings, RD, Esko, JD, Freeze, HH, Stanley, P, Bertozzi, CR, Hart, GW, Etzler, ME ed.), Cold Spring Harbor Laboratory Press, Cold Spring Harbor (NY). pp
10. Kunzler, M., Bleuler-Martinez, S., Butschi, A., Garbani, M., Luthy, P., Hengartner, M. O., and Aebi, M. (2010) *Methods Enzymol* **480**, 141-150
11. Kruger, R. P., Winter, H. C., Simonson-Leff, N., Stuckey, J. A., Goldstein, I. J., and Dixon, J. E. (2002) *J Biol Chem* **277**, 15002-15005
12. Ideo, H., Fukushima, K., Gengyo-Ando, K., Mitani, S., Dejima, K., Nomura, K., and Yamashita, K. (2009) *J Biol Chem* **284**, 26493-26501
13. Barrows, B. D., Griffitts, J. S., and Aroian, R. V. (2006) *Methods Enzymol* **417**, 340-358
14. Tateno, H., and Goldstein, I. J. (2004) *Arch Biochem Biophys* **427**, 101-109
15. Marroquin, L. D., Elyassnia, D., Griffitts, J. S., Feitelson, J. S., and Aroian, R. V. (2000) *Genetics* **155**, 1693-1699
16. Griffitts, J. S., Haslam, S. M., Yang, T., Garczynski, S. F., Mulloy, B., Morris, H., Cremer, P. S., Dell, A., Adang, M. J., and Aroian, R. V. (2005) *Science* **307**, 922-925

17. Barrows, B. D., Griffitts, J. S., and Aroian, R. V. (2007) *J Invertebr Pathol* **95**, 198-200
18. Gerdt, S., Dennis, R. D., Borgonie, G., Schnabel, R., and Geyer, R. (1999) *Eur J Biochem* **266**, 952-963
19. Suzuki, T. (2005) *Methods* **35**, 360-365
20. Colaert, N., Helsens, K., Martens, L., Vandekerckhove, J., and Gevaert, K. (2009) *Nat Methods* **6**, 786-787
21. Tompa, P., Buzder-Lantos, P., Tantos, A., Farkas, A., Szilagyi, A., Banoczi, Z., Hudecz, F., and Friedrich, P. (2004) *J Biol Chem* **279**, 20775-20785
22. Saido, T. C., Sorimachi, H., and Suzuki, K. (1994) *FASEB J* **8**, 814-822
23. Sorimachi, H., Ishiura, S., and Suzuki, K. (1997) *Biochem J* **328 (Pt 3)**, 721-732
24. Mo, H., Winter, H. C., and Goldstein, I. J. (2000) *J Biol Chem* **275**, 10623-10629
25. Ohm, R. A., de Jong, J. F., Lugones, L. G., Aerts, A., Kothe, E., Stajich, J. E., de Vries, R. P., Record, E., Levasseur, A., Baker, S. E., Bartholomew, K. A., Coutinho, P. M., Erdmann, S., Fowler, T. J., Gathman, A. C., Lombard, V., Henrissat, B., Knabe, N., Kues, U., Lilly, W. W., Lindquist, E., Lucas, S., Magnuson, J. K., Piumi, F., Raudaskoski, M., Salamov, A., Schmutz, J., Schwarze, F. W., vanKuyk, P. A., Horton, J. S., Grigoriev, I. V., and Wosten, H. A. (2010) *Nat Biotechnol* **28**, 957-963
26. Zhang, B., Palcic, M. M., Mo, H., Goldstein, I. J., and Hindsgaul, O. (2001) *Glycobiology* **11**, 141-147
27. Tateno, H., Winter, H. C., and Goldstein, I. J. (2004) *Biochem J* **382**, 667-675
28. Pechan, T., Ye, L., Chang, Y., Mitra, A., Lin, L., Davis, F. M., Williams, W. P., and Luthe, D. S. (2000) *Plant Cell* **12**, 1031-1040
29. Pechan, T., Cohen, A., Williams, W. P., and Luthe, D. S. (2002) *Proc Natl Acad Sci U S A* **99**, 13319-13323
30. Mohan, S., Ma, P. W., Pechan, T., Bassford, E. R., Williams, W. P., and Luthe, D. S. (2006) *J Insect Physiol* **52**, 21-28
31. van der Hoorn, R. A., and Jones, J. D. (2004) *Curr Opin Plant Biol* **7**, 400-407
32. Mohan, S., Ma, P. W., Williams, W. P., and Luthe, D. S. (2008) *PLoS One* **3**, e1786
33. McGhee, J. D. (2007) *WormBook*, 1-36
34. Lencer, W. I., and Saslowsky, D. (2005) *Biochim Biophys Acta* **1746**, 314-321
35. Beddoe, T., Paton, A. W., Le Nours, J., Rossjohn, J., and Paton, J. C. (2010) *Trends Biochem Sci* **35**, 411-418
36. Warner, R. L., Winter, H. C., Speyer, C. L., Varani, J., Oldstein, I. J., Murphy, H. S., and Johnson, K. J. (2004) *Exp Mol Pathol* **77**, 77-84
37. Kim, J. H., Johannes, L., Goud, B., Antony, C., Lingwood, C. A., Daneman, R., and Grinstein, S. (1998) *Proc Natl Acad Sci U S A* **95**, 2997-3002
38. Ong, D. S., Mu, T. W., Palmer, A. E., and Kelly, J. W. (2010) *Nat Chem Biol* **6**, 424-432
39. Halachmi, D., and Eilam, Y. (1989) *FEBS Lett* **256**, 55-61

ACKNOWLEDGMENTS

We thank I. J. Goldstein (Department of Biological Chemistry, University of Michigan Medical School, Ann Arbor, USA) and H. Tateno (National Institute of Advanced Industrial Science and Technology, Tsukuba, Japan) for providing plasmids MOApT7LO and MOA(Q46A,W138A)pT7LOH, R. Aroian (University of California, San Diego, USA) for providing plasmid pQE9(Cry5B), and Dr. P. Nanni (Functional Genomics Center, University and ETH Zürich, Switzerland) for support in mass spectrometry. This work was supported by the European Commission's Marie Curie programme (EuroGlycoArrays ITN), the Swiss National Science Foundation (grant no. 31003A-130671 to M. K., M. O. H. and M. A.) and ETH Zürich.

FIGURE LEGENDS

Figure 1. Recombinant MOA displays carbohydrate-binding dependent toxicity towards *C. elegans*. *E. coli* BL21(DE3) cells expressing MOA, the C-terminal deletion mutant MOA(ΔC), the carbohydrate-binding deficient variant MOA(Q46A,W138A) or containing empty vector were fed to *Caenorhabditis elegans* wild type N2 and *pmk1(km25)* mutant worms. *A.* MOA inhibits *C. elegans* development. The fraction of *C. elegans* L1 larvae having developed to L4 stage was scored after 72 h of feeding on a lawn of respective *E. coli* BL21 cells. Columns represent the average of 9 biological replicates. Error bars indicate standard deviations. *B.* MOA damages the *C. elegans* intestine. L4 larvae of *C. elegans* wild-type N2 were fed with MOA-expressing (lower panels) and control (upper panels) *E. coli* BL21(DE3) cells and examined after 24 h under the stereomicroscope (left panels) and by differential interference contrast (DIC) microscopy (right panels). Size bars, 20µm.

Figure 2. Toxicity of MOA towards *C. elegans* is mediated by binding to glycosphingolipids. *A.* Identified insertion sites of *Mos1* transposon in the *bre-3* gene of MOA-resistant *C. elegans* mutants. Arrows above *Mos1* elements indicate the orientation of the *Mos1* primer oJL115 used for sequencing iPCR products of mutant lysates. Bold letters indicate *C. elegans* genomic sequences followed by *Mos1* sequence. The gene model is taken from WormBase Release WS213. Clear boxes indicate exons and shaded boxes at the ends of the gene indicate 5'- and 3'-untranslated regions. *B.* Resistance of *C. elegans bre-3(ye26)*, *bre-4(ye27)* and *bre-5(ye17)* mutants towards MOA-mediated toxicity. *C. elegans* mutants of the indicated genotypes were analyzed for development from L1 to L4 as outlined above. Columns represent the average of 9 biological replicates. Error bars indicate standard deviations. *C.* MOA binds glycolipids of N2 and *bre-2(ye31)* but not *bre-4(ye27)* animals. Lipids were extracted from N2, *bre-2(ye31)* and *bre-4(ye27)* worms, resolved by TLC and stained with orcinol (lane 2-4) or overlaid with biotinylated Cry5B (lane 1) or MOA (lanes 5-7). The origin is always at the bottom. Letters refers to the glycolipid classification by Griffitts *et al.* (16). *D.* Tentative structure of the main glycosphingolipid species (component D) recognized by MOA. The core of the ceramide(Cer)-linked oligosaccharide is synthesized by BRE-3, BRE-5 and BRE-4. The terminal α1,3Gal is added by a yet unknown galactosyltransferase.

Figure 3. Purified recombinant MOA shows *in vitro* protease activity. *In vitro* protease activity of MOA using denatured RNase A as substrate. 0.5 (A, B) or 1 µg (C) of purified recombinant MOA and MOA(C215A) were incubated with 20 µg of denatured RNase A in assay buffer (40 mM Tris-HCl pH 8.0, 1 mM $CaCl_2$, 1% NP-40) for 1 h at 37 °C. The reaction was stopped by addition of SDS sample buffer and heating at 95 °C for 10 min. Samples were analyzed by SDS-PAGE and Coomassie blue staining. m, molecular weight standard. *A.* Catalytic activity of MOA is dependent on residue C215 and abolished by cysteine protease inhibitor E-64. MOA was incubated with 50 µM E-64, 2 mM PMSF and 2 mM $AgNO_3$ for 20 min at RT before the substrate was added. *B.* Ca^{2+} is required for protease activity of MOA. The reaction was performed in assay buffer w/o $CaCl_2$ as well as in buffer containing 1mM $CaCl_2$, $MgCl_2$, $MnCl_2$, and $ZnCl_2$ respectively. *C.* MOA protease activity has an alkaline pH optimum. Buffers at different pH values ranging from 4.5 to 9.7 were supplemented with 1mM $CaCl_2$ and 1% NP-40 and used as assay buffers. *D.* Cleavage site specificity of MOA. The iceLogo (20) was generated from 83 MOA-induced cleavage sites in native model substrates (bovine asialofetuin, bovine casein and human hemoglobin). MOA-mediated cleavage occurs between position 4 and 5 of the sequence logo.

Figure 4. Protease activity of MOA is required for full toxicity towards *C. elegans*. *A.* MOA(C215A) exhibits reduced toxicity towards *C. elegans*. *C. elegans* wild-type N2 and *pmk1(km25)* mutants were seeded onto lawns of *E. coli* BL21(DE3) cells expressing MOA, MOA(C215A) or containing empty vector and analyzed for development from L1 to L4 as described above. Columns represent the average of 9 biological replicates. Error bars indicate standard deviations. *B.* Cysteine protease inhibitor E-64 diminishes MOA-mediated toxicity towards *C. elegans*. L1 staged *C. elegans pmk-1(km25)* worms were seeded onto a lawn of *E. coli* BL21(DE3)

expressing MOA mixed with E-64 diluted in M9 buffer at the indicated concentrations. Development of the animals from L1 to L4 was analyzed as outlined above. As a control the experiment was performed with *E. coli* BL21(DE3) containing empty vector. Data shown represent the average of 9 biological replicates. Error bars indicate standard deviations.

Figure 1

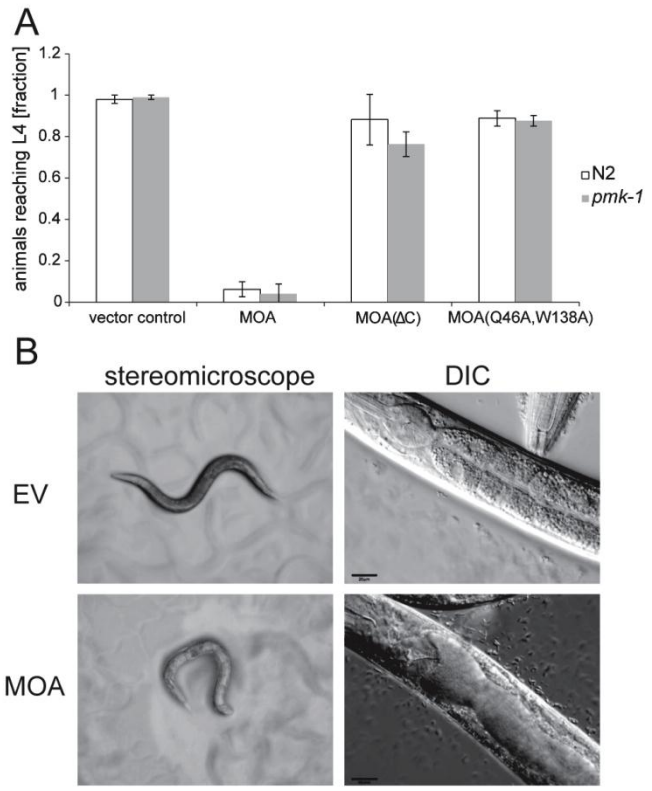

Figure 2

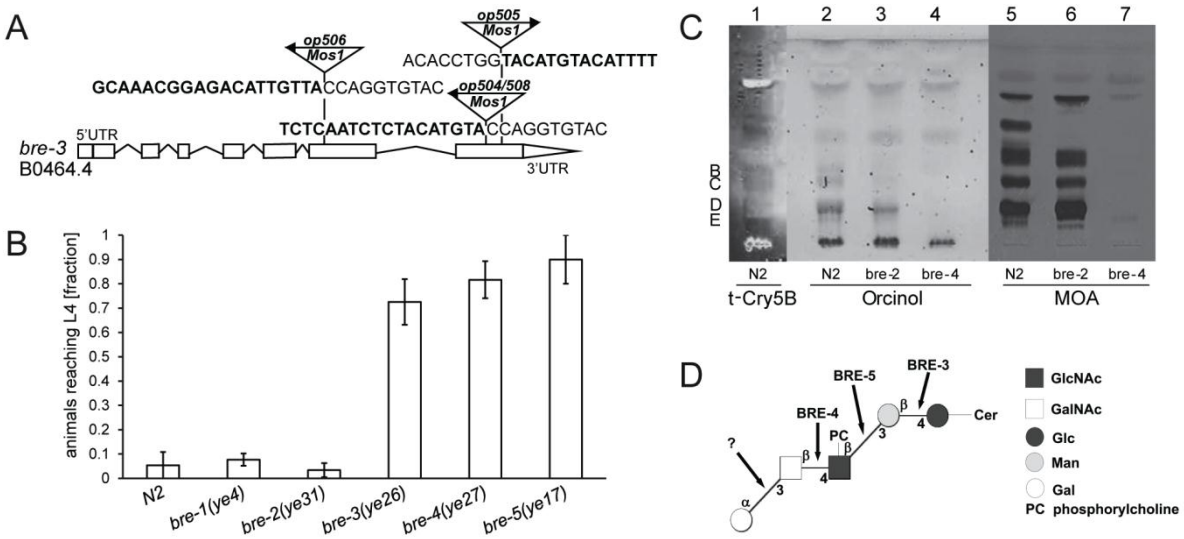

Figure 3

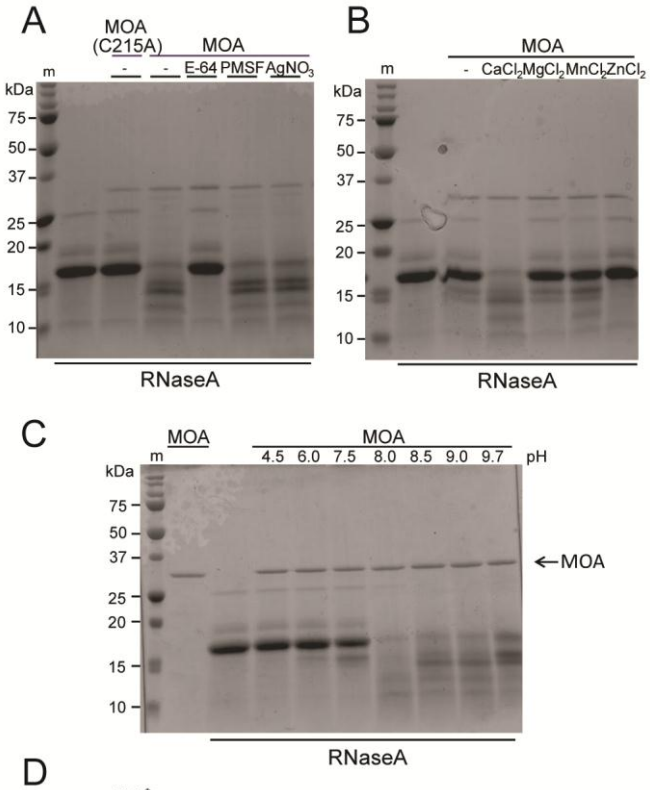

Figure 4